Advice-Efficient Prediction with Expert Advice

Yevgeny Seldin[1,2,3]   Peter Bartlett[1,2,4]   Koby Crammer[5]

[1]Mathematical Sciences School, Queensland University of Technology, Brisbane, QLD, Australia
[2]Department of Electical Engineering and Computer Sciences, UC Berkeley, Berkeley, CA, USA
[3]Department of Computer Science, University College London, London, UK
[4]Department of Statistics, UC Berkeley, Berkeley, CA, USA
[5]Department of Electrical Engineering, The Technion, Haifa, Israel

Abstract

Advice-efficient prediction with expert advice (in analogy to label-efficient prediction) is a variant of prediction with expert advice game, where on each round of the game we are allowed to ask for advice of a limited number M out of N experts. This setting is especially interesting when asking for advice of every expert on every round is expensive. We present an algorithm for advice-efficient prediction with expert advice that achieves $O\left(\sqrt{\frac{N}{M}T\ln N}\right)$ regret on T rounds of the game.

1 Motivation

We study the problem of prediction with expert advice in a setting, where we have a large set of experts N, but asking for advice of all experts on all rounds of the game is overly expensive. For example, the experts may be financial advisers designing investment portfolios for the stock market and getting the advice of each expert may be very expensive. Therefore, we would like to ask for advice of a smaller number $M \leq N$ of experts on each round (generally, $M \ll N$), but still be close to the best we could get if we would ask all experts for their advices. We call this setting advice-efficient prediction with expert advice in analogy to label-efficient prediction with expert advice (Cesa-Bianchi and Lugosi, 2006).

2 Setting and Notations

We work in prediction with expert advice setting (Cesa-Bianchi and Lugosi, 2006). We denote the action space by $\mathcal{X}$, the outcome space by $\mathcal{Y}$, and the loss function by $\ell : \mathcal{X} \times \mathcal{Y} \to [0, 1]$ (for our analysis there is no need to assume that the loss is convex in the first parameter). The number of experts is denoted by N and the experts are indexed by $h \in \{1, \ldots, N\}$. On each round i of the game each expert h produces an advice $\psi_i^h \in \mathcal{X}$. On each round the player is allowed to ask for advice of a fixed number $M \leq N$ of experts. The player asks for advice and plays action $X_i \in \mathcal{X}$. The environment then reveals an outcome y_i and the player suffers a loss $\ell(X_i, y_i)$ and the experts suffer losses $\ell(\psi_i^h, y_i)$. The goal of the algorithm is to minimize the regret defined as $\sum_{i=1}^{t} \ell(X_i, y_i) - \min_h \left(\sum_{i=1}^{t} \ell(\psi_i^h, y_i) \right)$.

3 Main Result

We prove the following regret bound for the algorithm presented in Algorithm 1 box.

Theorem 1. *The expected regret of Algorithm 1 on T rounds of the game satisfies:*

$$\mathbb{E}\left[\sum_{i=1}^{T} L_i \right] - \min_h \left(\sum_{i=1}^{T} \ell(\psi_i^h, y_i) \right) \leq 2\sqrt{\frac{N}{M}T \ln N}.$$

Algorithm 1: Advice-efficient prediction with expert advice.

$\forall h: \hat{L}_0(h) = 0$.

for $i = 1, 2, \ldots$ **do**

Let
$$q_i(h) = \frac{e^{-\eta_i \hat{L}_{i-1}(h)}}{\sum_{h'} e^{-\eta_i \hat{L}_{i-1}(h')}}.$$

Sample one expert H_i according to q_i. Get advice $\psi_i^{H_i}$.

Play $X_i = \psi_i^{H_i}$.

Observe nature outcome y_i and suffer loss $L_i = \ell(X_i, y_i)$.

Sample $M - 1$ additional experts uniformly without replacement. Let $\mathbb{1}_i^h = 1$ if expert h was sampled and $\mathbb{1}_i^h = 0$ otherwise. (For H_i used in the definition of X_i we have $\mathbb{1}_i^{H_i} = 1$.)

Get advices ψ_i^h for the experts sampled.

$$\forall h : L_i^h = \ell(\psi_i^h, y_i) \frac{1}{q_i(h) + (1 - q_i(h))\frac{M-1}{N-1}} \mathbb{1}_i^h.$$

$$\forall h : \hat{L}_i(h) = \sum_{j=1}^{i} L_j^h.$$

end

The "price" that we pay for observing the advice of M instead of all N experts is multiplicative $\sqrt{\frac{N}{M}}$ term. The constant is identical to the constant in the "simple" analysis of exponentially weighted forecasters in Cesa-Bianchi and Lugosi (2006, Corollary 2.2) and slightly worse than the constant in the tighter analysis in Cesa-Bianchi and Lugosi (2006, Theorem 2.3) (we are loosing a $\sqrt{2}$ factor), but we can improve the constant using similar techniques.

4 Analysis

The analysis is based on the following lemma, which follows from the analysis of EXP3 by Bubeck (2010).

Lemma 2. *For any N sequences of random variables $L_1^h, L_2^h, \ldots$ indexed by $h \in \{1, \ldots, N\}$, such that $L_i^h \geq 0$, and any non-increasing sequence $\eta_1, \eta_2, \ldots$, such that $\eta_i \geq 0$, for $q_i(h) = \frac{\exp(-\eta_i \sum_{j=1}^{i-1} L_j^h)}{\sum_{h'} \exp(-\eta_i \sum_{j=1}^{i-1} L_j^{h'})}$ (assuming for $i = 1$ the sum in the exponent is zero), for all $h^\star$ simultaneously we have:*

$$\sum_{i=1}^{T} \sum_{h} q_i(h) L_i^h \leq \sum_{i=1}^{T} \frac{\eta_i}{2} \sum_{h} q_i(h) \left(L_i^h\right)^2 + \frac{\ln N}{\eta_T} + \sum_{i=1}^{T} L_i^{h^\star}. \tag{1}$$

Now we are ready to prove Theorem 1.

Proof of Theorem 1. We study $\sum_h q_i(h) L_i^h$ and $\sum_h q_i(h) \left(L_i^h\right)^2$ for the case of our algorithm. We have:

$$\mathbb{E}\left[L_i^h\right] = \ell(\psi_i^h, y_i).$$

And we have:

$$\mathbb{E}\left[\sum_h q_i(h)L_i^h\right] = \sum_h q_i(h)\mathbb{E}\left[L_i^h\right] = \sum_h q_i(h)\ell(\psi_i^h, y_i) = \mathbb{E}\left[L_i\right]. \tag{2}$$

We also have:

$$\sum_h q_i(h)\left(L_i^h\right)^2 = \sum_h q_i(h)\left(\ell(\psi_i^h, y_i)\frac{1}{q_i(h) + (1-q_i(h))\frac{M-1}{N-1}}\mathbb{1}_i^h\right)^2$$

$$= \sum_h q_i(h)\ell(\psi_i^h, y_i)^2\left(\frac{1}{q_i(h) + (1-q_i(h))\frac{M-1}{N-1}}\right)^2\mathbb{1}_i^h$$

$$\leq \sum_h q_i(h)\left(\frac{1}{q_i(h) + (1-q_i(h))\frac{M-1}{N-1}}\right)^2\mathbb{1}_i^h.$$

And from here:

$$\mathbb{E}\left[\sum_h q_i(h)\left(L_i^h\right)^2\right] \leq \sum_h q_i(h)\left(\frac{1}{q_i(h) + (1-q_i(h))\frac{M-1}{N-1}}\right)^2\mathbb{E}\left[\mathbb{1}_i^h\right]$$

$$= \sum_h q_i(h)\frac{1}{q_i(h) + (1-q_i(h))\frac{M-1}{N-1}}$$

$$= \sum_h \frac{q_i(h)(N-1)}{q_i(h)(N-M) + M - 1}$$

$$\leq \frac{N}{M}. \tag{3}$$

The proof of the last inequality is provided in Lemma 3 the appendix.

By taking expectations of the two sides of (1) and substituting (2) and (3) we obtain for all $h^\star$:

$$\mathbb{E}\left[\sum_{i=1}^t L_i\right] \leq \frac{N}{M}\sum_{i=1}^t \frac{\eta_i}{2} + \frac{\ln N}{\eta_t} + \sum_{i=1}^t \ell(\psi_i^{h^\star}, y_i).$$

Finally, taking $\eta_i = \sqrt{\frac{M\ln N}{iN}}$ completes the proof. $\qquad\square$

5 Easy Extensions

The following extensions are easy to show:

1. Since the variance of L_i^h-s is bounded by $(N-1)/(M-1)$ independently of time, it is easy to derive a high-probability result with similar guarantees.

2. It is easy to show that the algorithm and analysis can be extended to adversarial multiarmed bandits, where we are allowed to reveal the loss of more than one action on each round (reward games can be translated to loss games via the transformation $\ell = 1 - r$, where $r \in [0,1]$ is the reward and $\ell \in [0,1]$ is the loss). Specifically, assume that in adversarial multiarmed bandit game with K arms the player plays and suffers the loss of one action on each round, but then the player is allowed to observe the losses of $M-1$ additional arms on the same round. Then, by identifying each arm with an expert that always predicts that arm, we can show that the regret of Algorithm 1 is $O\left(\sqrt{\frac{K}{M}T\ln K}\right)$. Interestingly, for $M > 1$ the variance of importance-weighted sampling is bounded by $(K-1)/(M-1)$ for all game rounds and it is possible to derive high-probability guarantees without additional smoothing in contrast to the EXP3.P algorithm.

References

Sébastien Bubeck. *Bandits Games and Clustering Foundations*. PhD thesis, Université Lille, 2010.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

A   Lemma 3

Lemma 3. *For any probability distribution q on $\{1, \ldots, N\}$ and any $M \leq N$:*

$$\sum_{h=1}^{N} \frac{q(h)(N-1)}{q(h)(N-M)+M-1} \leq \frac{N}{M}. \tag{4}$$

Proof. First, we show that the maximum of (4) is attained by the uniform distribution $q(h) = 1/N$. The Lagrangian corresponding to minimization of (4) subject to $\sum_h q(h) = 1$ is:

$$\mathcal{L}(q) = \sum_{h=1}^{N} \frac{q(h)(N-1)}{q(h)(N-M)+M-1} + \lambda \left(1 - \sum_h q(h)\right).$$

The first derivative of the Langrangian is:

$$\frac{\partial \mathcal{L}}{\partial q(h)} = \frac{(N-1)(q(h)(N-M)+M-1) - q(h)(N-M)(N-1)}{(q(h)(N-M)+M-1)^2} - \lambda = \frac{(N-1)(M-1)}{(q(h)(N-M)+M-1)^2} - \lambda.$$

The important point is that the derivative depends only on single h and, therefore, when we equate the derivative to zero the extremum is achieved when all $q(h)$ are equal. And, as a result, they are equal to $1/N$.

The second derivative is:

$$\frac{\partial^2 \mathcal{L}}{\partial q(h)^2} = -\frac{2(N-M)(N-1)(M-1)}{(q(h)(N-M)+M-1)^3} \leq 0$$

(note that for $M > 1$ and $N > M$ the inequality is strict; and for $M = 1$ or $N = M$ it is easy to check that (4) holds) and the mixed partial derivatives $\frac{\partial^2 \mathcal{L}}{\partial q(h)\partial q(h')} = 0$. Therefore, $q(h) = 1/N$ is the maximum point (for $1 < M < N$). Substituting $q(h) = 1/N$ into (4) we get:

$$\sum_{h=1}^{N} \frac{\frac{1}{N}(N-1)}{\frac{1}{N}(N-M)+M-1} = \frac{N(N-1)}{N-M+N(M-1)} = \frac{N(N-1)}{M(N-1)} = \frac{N}{M}.$$

$\square$